AN EXTENDED CONCEPT OF "MODEL"

E. S. Quade

July 1970

P-4427

14

AN EXTENDED CONCEPT OF "MODEL"

E. S. Quade
The Rand Corporation, Santa Monica, California

Operations research and systems analysis are now accepted, even extolled, as aids to management in commerce, industry, and defense. They are used not only to increase efficiency in routine operations but also to determine policies and actions, often at the highest levels. They have, however, not had, as yet, a corresponding acceptance for the solution of social and public problems.

One handicap may be a bias toward too narrow a concept of model.

Almost without exception, writers on operations research maintain that the major element in tackling a problem is to construct and use an operational model. "Its distinctive approach" as the Operational Research Society of Great Britain put it, "is to develop a scientific model of the system, incorporating measurements of factors such as chance and risk, with which to predict and compare the outcomes of alternative decisions, strategies or controls." An additional requirement is also specified; the model must represent or simulate the essential features of the situation under study. Such a model may take many forms, but the most useful, certainly the most used, and often

---

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The Rand Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The Rand Corporation as a courtesy to members of its staff.

A slightly abbreviated version of this paper was presented to the International Federation of Operational Research Societies at Venice, Italy, June 1969. It will appear in the Proceedings of the 5th International O.R. Conference, J. R. Lawrence (ed.), Tavistock Publications Limited, London, 1970.

the only types considered, are mathematical, expressing through a set of equations or a computer program the effectiveness and costs of alternative actions as a function of the variables one has under control. By operating with the model, either analytically or numerically, the consequences of alternative choices are determined and a preference identified.

A model in this tried but restricted sense is particularly difficult to create when political and social factors predominate. In industrial and military applications, the problem ordinarily deals with a man-made and directed enterprise--a manufacturing process, a weapon system, a transportation network. Goals can be defined, the design follows the laws of engineering and economics, and the essential relationships can be rediscovered and represented mathematically. In contrast, an attack on problems of air pollution, or urban renewal, or vocational rehabilitation, or criminal justice, involves working with goals that are obscure and conflicting and with a structure that may have grown without conscious design. To discover the underlying model may require the same sort of profound digging that is required to determine something like the role of hormones in regulating body functions. It is not surprising, therefore, that attempts to build quantitative models with which one can optimize in the conventional sense tend to fail. But this does not mean that operations research itself must also fail.

There are procedures for tackling these problems that promise more success. Unfortunately, they are not always accepted, or at least not much exploited, in operations research. To change this may require that we sharpen our conception of what a model contributes, if only to escape the view that it is simply a device to determine a preference by simulating reality.

The function of a model in operations research is "to predict and compare"; to provide a logical way to forecast the outcomes that follow alternative actions and, hopefully, to indicate a preference among them. A mathematical formulation with which one can optimize is an extremely valuable aid to this process. But it is not crucial; there are other routes. What is crucial to every decision process is reliance on expert judgment and intuition. This reliance permeates every aspect of operations--in deciding what approach is likely to be more fruitful, in designing the model, in determining what the facts are, and in interpreting the results. One great virtue of model building is that it provides a systematic, explicit, and efficient way to focus the required judgment and intuition.

A model, by introducing a precise framework and terminology, serves as an effective means of communication, enabling analysts and various experts* to exercise their judgment and intuition in a well-defined context and in proper relation to each other. In addition, it provides feedback to guide the participants in the revision of their earlier judgments. It is these features of the model that are essential to its role in supplying a route from hypotheses to prediction, not how explicitly it represents the real world or whether or not it provides a formal or quantitative scheme for optimization.

The realization that this is the case is neither new or startling. Operational gaming, that is to say, exercises in which the participants interact by playing roles that simulate individuals, or factions in a society, or even such things as sectors in an economy, is a step away from the traditional model and is now an accepted operations research technique. Its predictive quality, however, is very clearly a function of the intuitive insight provided by the participants. By allowing for the introduction of judgment at every step, a game provides an opportunity to

---

*A loose term, applied to anyone whose guidance and knowledge the analyst accepts.

take into account intangibles often considered completely beyond the reach of analysis. This is true both of the expert on the control team and of the player, who can let his decisions be influenced by his appraisal of the human effects of the simulated environment. For example, the success or failure of an economic plan may depend upon assumptions about a population's willingness to accept a change in diet or the flexibility of the political structure to accommodate a new power bloc. In any analytic formulation or computer simulation, factors of this type must be anticipated and decisions about them made in advance; in a game they can be made *seriatim*, in context, as the need arises.

But gaming--even though it sacrifices optimization--retains the representative features of the traditional model. My contention is that there are advantages in using approaches that sacrifice representation also. This suggests we take a broad view, accepting as a model *any device that provides a logical means to predict and compare the outcomes of alternative actions*, regardless of its representative features or how efficient it is at optimization. Calling such a device a "model" in the context of operations research would, I think, help to counter the bias toward mathematical models acquired by so many analysts through their education and work with industry.

To illustrate that there are real advantages in using models of this extended type, a discussion of one such device, the Delphi procedure, and an outline of how it might be used to tackle a problem with considerable social and political content, should suffice.

Delphi is an iterative procedure for eliciting and refining the opinions of a group of people by means of a series of questionnaires, a "framework" that replaces a representative model. In practice, the group would consist of experts or especially knowledgeable individuals, possibly including responsible decisionmakers. The idea is to

improve the panel or committee approach in arriving at a forecast or estimate by subjecting the views of the individual participants to each other's criticism in ways that avoid face-to-face confrontation. To this end, the process of deliberation is controlled, through feedback, by a steering group that preserves anonymity and computes the group response by using some form of averaging.

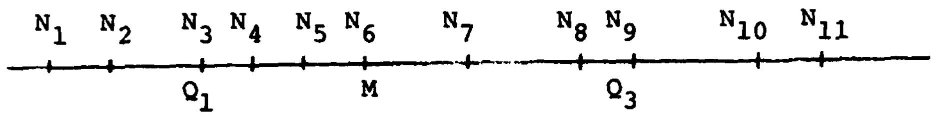
Anonymity serves to minimize the influence of vocal and persuasive individuals on group behavior. Also by having all interactions between respondents be through the steering group, "noise"--irrelevant or redundant material that obscures the directly relevant material offered by the participants--can be reduced.

The use of a statistical index, usually the median, to represent the group opinion is a further device to reduce group pressure toward conformity. No particular attempt to force unanimity among the respondents need then be made, and a spread of opinions on the final round is the normal outcome.

Let me now be more specific. Consider the common situation of having to arrive at an answer to the question of how large a particular number N should be. (For example, N could be the estimated cost of a measure, or a value representing its over-all benefit.) We might then proceed as follows:* (i) We would ask each participant independently to give an estimate of N, and then arrange the responses in order of magnitude, and determine the quartiles, Q_1, M, Q_3 so that the four intervals formed on the N-line by these three points each contain one quarter of the estimates. With eleven participants, the N-line might look like this:

---

*There are, of course, many possible variations.



(ii) We would communicate the values of Q_1, M, Q_3 to each respondent, ask him to reconsider his previous estimate and, if his new estimate lies outside the interquartile range (Q_1, Q_3), to state briefly the reason why, in his opinion, the answer should be lower (or higher) than that of the 75 percent majority opinion expressed in the first round. (iii) The results of this second round (which as a rule will be less dispersed than the first) would be communicated to the respondents in summary form, including the new quartiles and median. In addition, the reasons for raising or lowering the values, elicited in Round 2 and suitably collated and edited, would also be given to the respondents (always, of course, preserving anonymity). We would then ask the participants to consider the new estimates, giving the arguments the weight they deserve, and, in the light of the new information, to revise their previous estimates. Moreover, if a respondent's revised estimates fell outside the second round's interquartile range, he would be asked to state briefly why he found the argument that might have drawn his estimate toward the median unconvincing. (iv) Unless additional rounds seem advisable, the median of these Round 3 responses may then be taken as representing the group position as to what N should be.

Major credit for development of Delphi must be given to Olaf Helmer* and Norman Dalkey of Rand. My ideas

---
*Olaf Helmer is now at the Institute for the Future, Riverview Center, Middletown, Connecticut.

originated in their papers [1 - 6]. Although still experimental, Delphi has been used, among other tasks, to study educational innovations [7], to survey technological developments of interest to a commercial organization [8], and to provide short-range forecasts of business indices [9]. Except for the last, however, the superiority of such exercises over more standard methods is hard to assess.

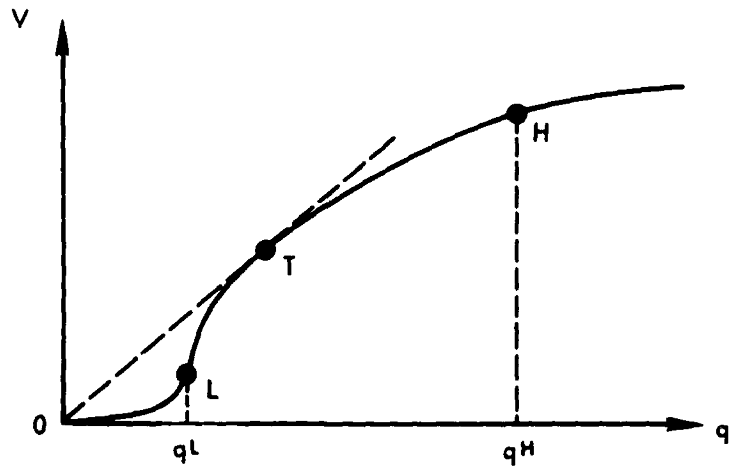
To illustrate the potentialities of Delphi for operations research, let me outline how it might be applied to a typical cost-effectiveness problem--allocating a budget for crime prevention.*

To begin, one might ask a panel drawn from the policymakers, their advisors, and experts familiar with the area to list projects that they feel should be included in any program. There will always be alternatives competing for funds: more police, better training, changes in court and parole procedure, new laws, and so forth. Not all promising projects can be financed; the problem is to devise a scheme to suggest and compare alternatives, and to select a preferred allocation of the freely disposable residue of the budget.

A project is rarely, if ever, of an all-or-nothing kind; that is, there is associated with it a degree to which it can be executed. The problem therefore is not merely to select the best projects but also to decide how much of each. In general, the value, V, that will be obtained from the application of a project, will typically appear as a function of its degree of adoption, q, in the form of an S-shaped curve, as in the figure below.

---

*For an extended illustration of another example, see [10], pp. 335-342.



A Delphi procedure can then be used to obtain a consensus opinion of the panel regarding two points on this curve: the value q^L below which adoption of the project would be pointless, and the value q^H above which the marginal benefits are so small as to make adoption wasteful. (One would expect many estimates of q^H to be zero, indicating total rejection of the project.) The unit of measurement of q is best in some natural unit, such as number of patrol cars; but if this is not possible some monetary unit may be used.

After reaching the best estimates of q^L and q^H, the next step is to ask a team of people with costing experience to work out and estimate the cost of these values of q for each project. (Of course the expected cost of a project depends to some extent on other projects that are adopted, but we must ignore this interaction at this stage.) Depending on the state of our knowledge about costing, we might or might not use a Delphi procedure here. Next, using these estimates of cost, we can construct a curve of cost versus quantity for each measure or project.

Finally, comparative estimates of effectiveness for each project must be obtained from the original panel. One scheme to do this might be to ask each panelist to allocate (using the cost-quantity relationships just described) the available budget among the various projects in the way that he thinks would give maximum effectiveness. The individual estimates can then be combined to give a group budget allocation.

This approach has many deficiencies. For one, were the budget to be implemented, it is unlikely that the actual costs and benefits would be identical with those determined separately for each project in isolation. Consequently, one or more iterations are called for so that the participants may take into account the existence of other projects at approximately the proper level. There may be more serious difficulties in this application, however. The results from any Delphi process, of course, depend on two critical factors: the choice of the panel and the way in which the process is implemented. Here, the choice of "policy advisors" and "experts in the area" for the panel may bias the outcome toward conservatism in dealing with a situation where the only hope for improvement may lie in innovation. And the topic is so shot through with judgments about values and goals that one may not like to trust the decision to such a panel, let alone to the group of analysts conducting the exercise.

Much remains to be learned about Delphi and the use of expertise. For example, we would like to know how much of the convergence that takes place is induced by the nature of the process itself rather than by elimination of the basic causes of disagreement. Placing the onus of justifying their responses on the respondents clearly tends to have the effect of causing those without strong convictions to move their estimates closer to the

median, for those who originally felt they had a good argument for a "deviationist" opinion may tend to give up their control estimate too easily; this may result in increasing the bandwagon effect instead of reducing it as intended.

One great drawback to written questionnaires is that the exchange is time consuming. To alleviate this, it is possible for the respondents to communicate with the steering group by typewriter or graphic consoles connected through an on-line computer system. Inputs and outputs can be in natural language.

The Delphi procedures--anonymous response, iteration, controlled feedback, numerical estimates, statistical "group response"--promise to become a highly effective means for group information processing. Experiments in which the respondents seek answers to "factual" questions pretty well indicate that:

1. Face-to-face discussion is not as efficient as the more formalized Delphi procedures.
2. Improvement in accuracy of estimates may be expected with an increase in the number of respondents, with iteration, and with estimates of range rather than simple point estimates.
3. The accuracy of estimates decreases as the time allowed for response increases and with discussion.

In addition, a certain amount of diffuse "experience" with Delphi suggests that the structural properties of the procedure leads to an enhanced acceptance of the group response by the individual participants beyond that obtained with more conventional (e.g., face-to-face) procedures. This is clearly a valuable characteristic, especially if the group is one of decisionmakers or others whose concurrence is required for the implementation of policy. It needs further study, however.

The potential usefulness of the Delphi approach is much wider than the published applications indicate [11, 12]. Extensive interest has been demonstrated by industrial and urban planners, research managers, and policymakers (in the U.S. government and elsewhere) in the promise of Delphi procedures for technological forecasting, corporate planning, organizational decisionmaking, and policy evaluation. Suggested applications range from the drafting of diplomatic notes and long-range political forecasting to determining what products to market. Unfortunately, many of the applications being considered are marginal at the moment, in the sense that greater effectiveness of Delphi procedures over more conventional techniques has yet to be demonstrated.

In view of the accelerating interest in the use of Delphi procedures by such a wide spectrum of public and private institutions, two topics, in addition to that suggested in this paper, are of immediate practical concern: its use for forecasting technological and social events and for making value judgments.

With regard to forecasting, experimentation with short-range predictions suggests that the conclusions from factual estimation experiments apply to them as well; but this presumption needs confirmation. Not much has been done experimentally to verify the value of Delphi for long-range forecasts in comparison with other methods, but experiments to investigate the reliability of such forecasts, in the technical sense of consistency of judgments by similar groups of "experts," are being considered.

In the area of value judgments, as with long-range predictions, there is not too much that can be done with regard to "accuracy," but the reliability and stability of group evaluation are being investigated experimentally. There is evidence from applied studies that iteration produces convergence with value judgments, but whether this convergence is stable or capricious is not visible from these uncontrolled exercises.

But imperfect as it is, the Delphi process or some further modification appears to promise much for the investigation of problems with a high social and political content. Because it can be used to allocate resources rationally and to force explicit thinking about the measurement of benefits, Delphi offers a hope of introducing cost-effectiveness thinking into a wide range of problems where conventional models are difficult to formulate. If Delphi is not already considered to be an operations research procedure, we should extend our concept of "model" enough to bring it within the fold.

REFERENCES

1. Helmer, O., Social Technology, Basic Books, 1966.

2. Helmer, O., and N. Rescher, "On the Epistomology of the Inexact Sciences," Management Science, 6, 1959, p. 47.

3. Dalkey, N., and O. Helmer, "An Experimental Application of the Delphi Method to the Use of Experts," Management Science, 9, 1963, p. 458.

4. Dalkey, N., Delphi, The Rand Corporation, P-3704, October 1967.

5. -----, "An Experimental Study of Group Opinion," Futures, September 1969, pp. 408-426.

6. -----, "Analysis from a Group Opinion Study," Futures, December 1969, p.p. 541-551.

7. Adelson, Marvin (ed.), "Planning Education for the Future," American Behavioral Scientist, Vol. 10, No. 7, March 1967, pp. 1-31.

8. North, H. Q., A Probe of TRW's Future: The Next 20 Years, TRW Systems, Redondo Beach, California, July 1966. See also, "Setting a Timetable," Business Week, Issue 1969, May 27, 1967, pp. 52-61.

9. Campbell, Robert H., A Methodological Study of Expert Utilization in Business Forecasting, unpublished doctoral dissertation, University of California, Los Angeles, 1966.

10. Quade, E. S., and W. I. Boucher, Systems Analysis and Policy Planning: Applications in Defense, American Elsevier, New York, 1968, Chapter 18.

11. Helmer, O., T. J. Gordon, S. Enzer, R. deBrigard, and R. Rochberg, Development of Long-Range Forecasting Methods for Connecticut: A Summary, R-5, Institute for the Future, Middletown, Connecticut, September 1969.

12. Gordon, T. J., and R. H. Ament, Forecast of Some Technological and Scientific Developments and Their Societal Consequences, R-6, The Institute for the Future, Middletown, Connecticut, September 1969.